Research Institute
Social Cohesion

RDC

SOEP

SCP Documentation

German Social Cohesion Panel

SCP 2021-23 W1-3 Codebook PBRUTTO: Person-Related Gross File (English)



Research Institute Social Cohesion

SOEP The Socio-Economic Panel

German Social Cohesion Panel

Established in 2021, the German Social Cohesion Panel (SCP) is a wide-ranging representative longitudinal study of private households in Germany, carried out in collaboration of the Research Institute Social Cohesion (RISC) and the German Socio-Economic Panel (SOEP).

The aim of the SCP Documentation is to thoroughly document the survey's data collection and data processing.

Recommended Citation

Groh-Samberg, O., Axenfeld, J. B., Gerlitz, J.-Y., Cornesse, C., Kroh, M., Lengfeld, H., Reinecke, J., Teichler, N., Traunmüller, R., & Zinn, S. (2026). SCP 2021-23 W1-3 - Codebook PBRUTTO: Person-Related Gross File (English). *German Social Cohesion Panel 2021-2023 - Wave 1-3*. Bremen and Berlin: RDC-RISC/SOCIUM, SOEP/DIW Berlin. doi:10.60532/scp.2021-23.w1-3.v1

- **Authors:** Olaf Groh-Samberg, Julian B. Axenfeld, Jean-Yves Gerlitz, Carina Cornesse, Martin Kroh, Holger Lengfeld, Jost Reinecke, Nils Teichler, Richard Traunmüller, Sabine Zinn

- **Contributors:** Cosima Adams, Anton Bochert, Martin Gerike, Josefine Kuhrmeier, Anna-Tabea Müller, Eric Nissen, Sebastian Rueda-Uribe, Rainer Siegers, Hans Walter Steinhauer, Knut Wenzig, Julia Witton (Project Members), infas (Data Collector)

- **Publisher:** RDC-RISC
  SOCIUM, University of Bremen
  P.O. Box 330 440
  28334 Bremen
  Germany

  SOEP
  DIW Berlin
  German Socio-Economic Panel (SOEP)
  Anton-Wilhelm-Amo-Str. 58
  10117 Berlin
  Germany

- **DOI:** 10.60532/scp.2021-23.w1-3.v1

- **Website:** www.fgz-risc-data.de

  www.diw.de

The text of this publication is published under the Creative Commons license CC BY-SA 4.0 Attribution-ShareAlike 4.0 International. The exact wording of the license CC BY-SA 4.0 can be found here:

https://creativecommons.org/licenses/by-sa/4.0/

SCP Documentation

German Social Cohesion Panel

SCP 2021-23 W1-3 Codebook PBRUTTO: Person-Related Gross File (English)

Contents

1 General Information **2**

2 Identifiers **2**
pid – Never Changing Person ID 2
hid – Current Household ID 3
cid – Original Household ID 3

3 Survey Context **4**
wave – Survey Wave 4
restype – Type of respondent 5

4 Demographics **5**
paddressdif – Different postal addresses in the HH 5
geburt_reg – Year of Birth 6
sex_reg – Sex 6
pnat_reg – Nationality 7

5 Participation status **7**
lint_scp – Wave Last Interviewed 7
paapor_scp – AAPOR Class. Person 7

1 General Information

The PBRUTTO dataset contains information about the individuals in the gross sample which is not asked in the individual questionnaire but is instead background information from the sampling frame or survey fieldwork. Variables that only contain information from the residents' registration offices have the suffix "_reg". The information from the residents' registration offices may differ from the survey data.
In some places in the documentation and in the data, year numbers are used, for example, for the names of variables and of the questionnaire instrument. These year numbers are always based on the field start of the data collection of the corresponding survey wave.
Data in this dataset may be missing due to different reasons. Missings are coded "[-1] No answer" if the registration office fails to provide valid information that is required for the variable. Missings are coded "[-2] Does not apply" if the information is missing due to the study design. Information from the residents' registration offices is only available for the initial anchor persons.

2 Identifiers

pid – Never Changing Person ID

2110000101		1
2110000201		1
2110000301		4
2110000302		1
2110000401		1
2110000501		1
2110000601		1
2110000701		1
2110000801		1
2110000901		4
2110000902		4
2110001001		1
2110001002		1
2110001101		1
2110001201		1
...	(44554 rows omitted)	79507
2113798801		1
2113798901		1
2113799001		1
2113799101		4
2113799102		4
2113799201		1
2113799301		1
2113799401		1
2113799501		1
2113799601		1
2113799701		1
2113799801		1
2113799901		1
2113800001		1

2210389603	1

The central individual identifier across time is PID, which is fixed over time (and of course datasets).

hid – Current Household ID

21100001		1
21100002		1
21100003		5
21100004		1
21100005		1
21100006		1
21100007		1
21100008		1
21100009		8
21100010		2
21100011		1
21100012		1
21100013		1
21100014		1
21100015		1
...	(37902 rows omitted)	79510
22250038		1
22250039		1
22250040		1
22250041		1
22250042		1
22250043		1
22250044		1
22250045		1
22250046		1
22250047		1
22250048		1
22250049		1
22250050		1
22250051		1
22250052		1

This identifier groups all individuals into their respective households at the time of the most recent wave. This means that a person's HID can change over time, for example if an adult child moves out of their parents' home and starts their own household. The HID reflects the status prior to the fieldwork phase. Changes in household membership are only accounted for in the HID of the subsequent wave.

cid – Original Household ID

21100001		1
21100002		1
21100003		5
21100004		1
21100005		1
21100006		1
21100007		1
21100008		1
21100009		8
21100010		2
21100011		1
21100012		1
21100013		1
21100014		1
21100015		1
...	(37844 rows omitted)	79500
21137986		1
21137987		4
21137988		1
21137989		1
21137990		1
21137991		8
21137992		1
21137993		1
21137994		1
21137995		1
21137996		1
21137997		1
21137998		1
21137999		1
21138000		1

This identifier groups individuals into their original households at the start of the panel. That means that a person's CID is time-constant and will always relate them back to the household they initially belonged to, even if they moved out since.

3 Survey Context

wave – Survey Wave

1	[1] Wave 1, part 1 (2021/22)	43819
2	[2] Wave 1, part 2 (2021/22)	11654
3	[3] Wave 2 (2022/23)	12079
4	[4] Wave 3 (2023)	12000

This variable identifies the (partial) wave in which the data collection took place.

restype – Type of respondent

1	[1] Anchor person (AP)	64236
2	[2] Household member (HM)	14551
3	[3] New household member (NHM)	765

Anchor persons (AP) are respondents at the time of recruitment who were drawn as part of the initial random sample at the residents' registration offices and invited to take part in the survey. In addition to the personal questionnaire, these people usually also complete the household questionnaire. Household members (HM) are respondents who were specified by anchor persons as living in the same household aged 18 and over and were also invited to take part in the survey. HM only complete the personal questionnaire. New household members (NHM) are respondents who were additionally named as household members by anchors from wave 2 onwards. In particular, this includes people who moved into the household or who have reached the age of 18 in the period since the last survey. These persons receive an extended personal questionnaire and, with their consent, are integrated into the panel as HM from the following wave onwards.
If HM or NHM leave the household after inclusion in the panel, these individuals themselves become the AP of the new household in the following wave, unless this role is filled by another member of the household. Similarly, a HM or NHM becomes the new AP if the original AP leaves the household or refuses further interviews.

4 Demographics

paddressdif – Different postal addresses in the HH

0	[0] Postal address matches that of the AP	22929
1	[1] 1st different postal address within the HH	310
2	[2] 2nd different postal address within the HH	10
999	[999] Incomplete or incorrect address details	830
-1	[-1] No answer	0
-2	[-2] Does not apply	0
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the questionnaires this wave	55473
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

This variable identifies individuals whose reported postal address at the end of the survey period differs from that of their anchor person. This is usually the result of either at least one person moving out of the shared household, or the household moving house whilst one person did not participate in the relevant wave. In the first case, a split household is usually formed in the subsequent wave, with one of the participating household members being designated as the reference person for this new household. In the second case, the address

of the participating person(s) is usually adopted in the subsequent wave for those who did not participate in the most recent wave as well.
The address coordinates of the participants are used to generate the variable. The variable takes the value 0 if a person's reported postal address matches that of the anchor person. If the postal address differs from that of the anchor person, different postal addresses within a household are numbered consecutively: persons with the same first differing postal address are assigned the value 1, persons with the same second differing postal address the value 2, and so on. If persons could not be unambiguously assigned to an address due to a missing, incorrect or incomplete address, they are coded as 999.

geburt_reg – Year of Birth

1921	[1921] 1921 and before	17
1922	[1922] 1922	6
1923	[1923] 1923	13
1924	[1924] 1924	20
1925	[1925] 1925	25
1926	[1926] 1926	39
1927	[1927] 1927	65
1928	[1928] 1928	93
1929	[1929] 1929	85
1930	[1930] 1930	109
1931	[1931] 1931	154
1932	[1932] 1932	191
1933	[1933] 1933	173
1934	[1934] 1934	308
1935	[1935] 1935	323
...	(64 rows omitted)	57974
2000	[2000] 2000	841
2001	[2001] 2001	855
2002	[2002] 2002	743
2003	[2003] 2003	586
2004	[2004] 2004	73
-1	[-1] No answer	661
-2	[-2] Does not apply	16198
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the questionnaires this wave	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

Birth years 1921 and earlier were combined into one category for data protection reasons.

sex_reg – Sex

1	[1] Male	31511
2	[2] Female	31833
3	[3] Diverse	0
4	[4] No entry (official designation)	0
-1	[-1] No answer	10
-2	[-2] Does not apply	16198
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the questionnaires this wave	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

pnat_reg – Nationality

1	[1] German	54395
2	[2] Not German	5997
-1	[-1] No answer	2962
-2	[-2] Does not apply	16198
-3	[-3] Implausible value	0
-4	[-4] Inadmissible multiple response	0
-5	[-5] Not included in this version of the questionnaire	0
-6	[-6] Version of questionnaire with modified filtering	0
-7	[-7] Only available in less restricted edition	0
-8	[-8] Question not part of the questionnaires this wave	0
-9	[-9] Don't want to answer [only CAWI]	0
-10	[-10] Break-off [only CAWI]	0

5 Participation status

lint_scp – Wave Last Interviewed

0	[0] No Interview	27122
1	[1] Wave 1, part 1 (2021/22)	23027
2	[2] Wave 1, part 2 (2021/22)	11426
3	[3] Wave 2 (2022/23)	9985
4	[4] Wave 3 (2023)	7992

paapor_scp – AAPOR Class. Person

1100	[1100] 1.1 Complete	41254
1200	[1200] 1.2 Partial	1135
2111	[2111] 2.111 Refusal (this wave)	178

2112	[2112] 2.112 Known respondent refusal	4
2120	[2120] 2.12 Break-off	440
2200	[2200] 2.2 Non-contact	31924
2310	[2310] 2.31 Dead	166
2321	[2321] 2.321 Temporary physically or mentally unable/incompetent (P Level)	3
2322	[2322] 2.322 Permanently physically or mentally unable/incompetent (P Level)	100
2331	[2331] 2.331 Language problem	8
2351	[2351] 2.351 Soft refusal (currently not willing no time)	928
2400	[2400] 2.4 New address after field period	288
3180	[3180] 3.18 Unable to locate address	804
3320	[3320] 3.32 Post: Moved left no address	1962
3400	[3400] 3.40 Returned with forwarding inform.	345
4200	[4200] 4.2 Moved abroad	13

Corresponds to the AAPOR standard definitions: https://aapor.org/standards-and-ethics/standard-definitions